REVSTAT – Statistical Journal
Volume 3, Number 1, June 2005, 19–44

ASYMPTOTIC BEHAVIOUR OF REGULAR ESTIMATORS

Authors: Jean Diebolt
– CNRS, Université de Marne-la-Vallée,
Équipe d'Analyse et de Mathématiques Appliquées,
5, boulevard Descartes, Batiment Copernic, Champs-sur-Marne,
77454 Marne-la-Vallée Cedex 2 (jean.dieb@wanadoo.fr)

Armelle Guillou
– Université Paris VI, Laboratoire de Statistique Théorique et Appliquée,
Boîte 158, 175 rue du Chevaleret
75013 Paris (guillou@ccr.jussieu.fr)

Received: November 2003 Revised: October 2004 Accepted: December 2004

Abstract:

- The P.O.T. (Peaks-Over-Threshold) approach consists of using the generalized Pareto distribution (GPD) to approximate the distribution of excesses over thresholds. We use the maximum likelihood estimators, or some other ones satisfying regularity conditions, to estimate the parameters of the approximating distribution. We study the asymptotic bias of these estimators and also the functional bias of the estimated GPD.

Key-Words:

- *Extreme values; domain of attraction; excesses; generalized Pareto distribution; maximum likelihood estimators.*

AMS Subject Classification:

- 60G70, 62G20.

1. INTRODUCTION

In many statistical applications it is necessary to make inferences about the tail of a distribution, where little data is available. For instance, one is often interested in the probability that the maximum of n random variables exceeds a given threshold or, *vice versa*, one wants to determine a level such that the exceedance probability is below a given small value. As an example, an hydraulics engineer has to estimate the necessary height of a dike such that the probability of a flooding in a given year is less than 10^{-4} (cf. Dekkers and de Haan, 1989). This interest has given rise to a rapid development of extreme value theory in the last thirty years (see e.g. Galambos, 1978, Leadbetter *et al.*, 1983). The traditional approach to the analysis of extreme values in a given population is based on the family of generalized extreme value (GEV) distributions. More precisely, Gnedenko (1943) showed that the limit distribution of the maximum $X_{n,n}$ of a sample of independent and identically distributed variables $X, X_1, ..., X_n$ from a distribution F, properly centred and normalized, is necessarily of extreme value type, i.e. for some $\gamma \in \mathbb{R}$, there exists sequences of constants $\sigma_n > 0$ and $\alpha_n \in \mathbb{R}$ such that

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{X_{n,n} - \alpha_n}{\sigma_n} \le x\right) \;\longrightarrow\; H_\gamma(x) \;, \tag{1.1}$$

for all continuity points of the extreme value distribution function H_γ, defined as

$$H_\gamma(x) \;=\; \begin{cases} \exp\bigl(-\,(1+\gamma x)^{-\frac{1}{\gamma}}\bigr) & \text{for} \quad \gamma \neq 0 \;\text{ and }\; 1+\gamma x > 0\,, \\ \exp\bigl(-\exp(-x)\bigr) & \text{for} \quad \gamma = 0 \;\text{ and }\; x \in \mathbb{R}\;. \end{cases}$$

The distribution function F is said to belong to the maximum domain of attraction of H_γ. The real-valued parameter γ is referred to as the extreme value index (EVI) of F. Most common continuous distribution functions satisfy this weak condition. Distributions with $\gamma > 0$ are called heavy-tailed, as their tail $\bar{F}$ typically decays slowly as a power function. Examples in this Fréchet class are the Pareto, Burr, Student's t, α-stable ($\alpha < 2$) and loggamma distributions. The Gumbel class of distributions with $\gamma = 0$ encompasses the exponential, normal, lognormal, gamma and classical Weibull distributions, the tail of which diminishes exponentially fast. Finally, the Weibull class consists of distributions with $\gamma < 0$, which all have a finite right endpoint $s_+(F) := \sup\{x : F(x) < 1\}$. Examples in this class are the uniform and reverse Burr distributions.

The problem of estimating the so-called extreme value index γ, which determines the behaviour of the underlying distribution function F in its upper tail, has received much attention in the literature. An extensive motivation of this estimation problem can be found in Galambos (1978). The GEV distribution is appropriate when the data consist of a set of maxima. However, there has been some criticism of this approach, because using only maxima leads to the loss of information contained in other large sample values in a given period. This problem is remedied by considering several of the largest order statistics instead of just the largest one: that is, considering all values larger than a given threshold. The differences between these values and a given threshold are called excesses over the threshold. Denote by $F_u(x) := \mathbb{P}(X - u \le x \mid X > u)$ the distribution of the excesses of X over u, given that u is exceeded, and by $G_{\gamma,\sigma}$ the generalized Pareto distribution (GPD) defined, for all $x \ge 0$, as

$$G_{\gamma,\sigma}(x) = \begin{cases} 1-\left(1+\dfrac{\gamma x}{\sigma}\right)^{-\frac{1}{\gamma}} & \text{for} \quad \gamma \neq 0 \ \text{ and } \ 1+\gamma x/\sigma > 0\,, \\ 1-\exp\left(-\dfrac{x}{\sigma}\right) & \text{for} \quad \gamma = 0\ , \end{cases}$$

where σ and γ are the scale and shape parameters.

Pickands' and Balkema and de Haan's result (see Pickands (1975) and Balkema and de Haan (1974)) on the limiting distribution of excesses over a high threshold states that condition (1.1) holds if and only if

$$\lim_{u \to s_+(F)} \sup_{0<x<s_+(F)-u} \left|F_u(x) - G_{\gamma,\sigma(u)}(x)\right| = 0$$

for some positive scaling function $\sigma(u)$ depending on u.

Thus, if, for some n, one fixes a high threshold u_n and selects from a sample $X_1, ..., X_n$ only those observations $X_{i_1}, ..., X_{i_{N_n}}$ that exceed u_n, a GPD with parameters γ and $\sigma_n = \sigma(u_n)$ is likely to be a good approximation for the distribution F_{u_n} of the N_n excesses $Y_j := X_{i_j} - u_n$, $j = 1, ..., N_n$. This is called the Peaks-Over-Threshold (P.O.T.) method.

Several methods have been proposed for estimating the parameters of the GPD, for example the method of moments, of the probability-weighted moments introduced by Hosking and Wallis (1987) or the maximum likelihood method (Smith, 1987). In this paper, we look in more details at the maximum likelihood estimators, but also we derive the more general conditions required on the estimators $(\hat{\gamma}_n, \hat{\sigma}_n)$ in order to obtain our results.

In all the sequel, we denote by

$$\mathbb{A}_{n,u_n}(x) = \frac{1}{N_n} \sum_{j=1}^{N_n} 1\!\!1_{\{Y_j \le x\}}$$

the empirical distribution function of the excesses.

It is of course very important to measure the error between $\bar{F}_{u_n} := 1 - F_{u_n}$ (unknown) and its estimator $\bar{G}_{\hat{\gamma}_n,\hat{\sigma}_n} := 1 - G_{\hat{\gamma}_n,\hat{\sigma}_n}$. This error can be splitted into two parts: an approximation error and an estimation error. The first one, also called *bias of approximation*, is justified by the fact that the distribution of the excesses over u_n is only approximated by a GPD, which implies a systematic error studied in Worms (2000). Since the distribution of the excesses over u_n is F_{u_n} and not G_{γ,σ_n}, the error due to the estimation of (γ,σ_n) is also divided into an approximation error due to the bias of approximation and a random term due to fluctuations.

Note that N_n follows a binomial distribution $\mathcal{B}(n,1-F(u_n))$. We suppose, in all the sequel, that $n(1-F(u_n)) \to \infty$ as $n \to \infty$, that is $N_n \to \infty$ in probability. In such a case, $\frac{N_n}{n(1-F(u_n))} \to 1$ in probability, as $n \to \infty$.

Let

$$F^*_{u_n}(y) = F_{u_n}(\sigma_n y) \qquad \text{and} \qquad \mathbb{A}^*_{n,u_n}(y) = \mathbb{A}_{n,u_n}(\sigma_n y)$$

for all $y \in \mathbb{R}_+$. We will study the asymptotic behaviour of

$$\bar{F}_{u_n}(x) - \bar{G}_{\hat{\gamma}_n,\hat{\sigma}_n}(x)$$

where $\hat{\gamma}_n$ and $\hat{\sigma}_n$ are the maximum likelihood estimators, or other regular estimators with properties specified later on.

In what follows, we suppose that F is twice differentiable and that its inverse F^{-1} exists. Let V and A be two functions defined as

$$V(t) = \bar{F}^{-1}(e^{-t}) \qquad \text{and} \qquad A(t) = \frac{V''(\ln t)}{V'(\ln t)} - \gamma \ .$$

We suppose the following first and second order conditions:

(1.2) $$\lim_{t\to+\infty} A(t) = 0 \ ,$$

and

(1.3) A is of constant sign at ∞ and there exists $\rho \le 0$ such that $|A| \in RV_\rho$,

(see Bingham *et al.*, 1987).

Under these assumptions, it is proved in Worms (2000) (Theorem 1.4, p. 43) that as $u_n \to s_+(F)$

(1.4) $$\bar{F}_{u_n}(\sigma_n y) - \bar{G}_\gamma(y) \ = \ a_n \, D_{\gamma,\rho}(y) + o(a_n), \qquad \text{as} \ \ n \to +\infty \ ,$$

for all y, when

$$\sigma_n := \sigma(u_n) = V'\big(V^{-1}(u_n)\big) \ , \qquad a_n := A\big(e^{V^{-1}(u_n)}\big)$$

$$\bar{G}_\gamma(y) := 1 - G_{\gamma,1}(y) \ ,$$

and

$$D_{\gamma,\rho}(y) \ = \ \begin{cases} C_{0,\rho}(y), & \text{if} \quad \gamma = 0 \, , \\[2ex] C_{\gamma,\rho}\Big(\dfrac{1}{\gamma}\ \ln(1+\gamma y)\Big) & \text{if} \quad \gamma \ne 0 \ , \end{cases}$$

where

$$C_{\gamma,\rho}(y) := \mathrm{e}^{-(1+\gamma)y}\, I_{\gamma,\rho}(y) \qquad \text{and} \qquad I_{\gamma,\rho}(y) := \int_0^y \mathrm{e}^{\gamma u} \int_0^u \mathrm{e}^{\rho s}\, ds\, du \;.$$

We also assume that

$$\lim_{n\to\infty} \sqrt{n\,(1-F(u_n))}\; a_n \;=\; \lambda \;\in\; \mathbb{R} \;. \tag{1.5}$$

This is equivalent to suppose that $\sqrt{N_n}\, a_n$ tends to λ in probability, as $n \to \infty$.

The main result of this paper is the following. For a regular class of estimators, when ρ is equal to 0, the error due to the fact that $\bar{F}_{u_n}$ is replaced by $\bar{G}_{\hat{\gamma}_n,\hat{\sigma}_n}$ is of smaller order than the same error in the case $\rho \neq 0$. This result is closely linked to the penultimate approximation for the distribution of the excesses established in Worms (2002) (Gomes and de Haan (2000), generalizing Cohen (1982), Gomes (1984) and Gomes and Pestana (1987), studied penultimate approximation for the distribution of the maximum). At first sight, it can appear a bit strange since it is well known that, if we consider only the problem of the estimation of the index γ, the smaller $|\rho|$, the more difficult it is to estimate the index. This problem of bias in the estimation of the index has been widely studied recently in the literature and justified in particular the work on regression model by Beirlant *et al.* (1999). This paper proves that, on the contrary, if we consider the problem of the estimation of the tail distribution, we do not need to construct asymptotically unbiased estimators, which is essential in the other estimation problem.

The remainder of our paper is organized as follows. In Section 2, we give our main results and the general conditions on the estimators $(\hat{\gamma}_n, \hat{\sigma}_n)$ that we need to obtain our results. Then, in Section 3, we study the asymptotic bias and also the functional bias of the estimated GPD in the case of maximum likelihood estimation with $\gamma > 0$. The details of the proofs are postponed to the appendix.

2. MAIN RESULT

We restrict our attention to a class of estimators that we call "regular estimators" of the couple (γ, σ). We say that an estimator is regular if it has the form $T(\bar{\mathbb{A}}_{n,u_n}) =: (T_1(\bar{\mathbb{A}}_{n,u_n}), T_2(\bar{\mathbb{A}}_{n,u_n}))$, where T satisfies:

$(\mathbf{A}_1)$ $T(\bar{G}_{\gamma,\sigma}) = \big(T_1(\bar{G}_{\gamma,\sigma}), T_2(\bar{G}_{\gamma,\sigma})\big) = (\gamma, \sigma)$.

$(\mathbf{A}_2)$ A form of Hadamard differentiability, namely the existence of linear forms $DT(\bar{G}_\gamma) =: (DT_1(\bar{G}_\gamma), DT_2(\bar{G}_\gamma))$, where

$$DT_k(\bar{G}_\gamma)[H] \;=\; \int_0^\infty H\, d\mu_{k,\gamma} \;, \qquad k = 1, 2 \;, \tag{2.1}$$

for some Borelian measures $\mu_{k,\gamma}$ and all $H \in L^1(\mathbb{R}^+, \mu_{1,\gamma}) \cap L^1(\mathbb{R}^+, \mu_{2,\gamma})$ such that under assumption (1.5) we have

$$\lim_{n\to+\infty} \frac{T(\bar F^*_{u_n}) - T(\bar G_\gamma)}{a_n} = DT(\bar G_\gamma)[D_{\gamma,\rho}] \,, \tag{2.2}$$

$$\lim_{n\to+\infty} \sqrt{N_n}\left(T\left(\bar F^*_{u_n} + \frac{1}{\sqrt{N_n}}\,\alpha_{N_n}\circ \bar F^*_{u_n}\right) - T(\bar F^*_{u_n})\right) = DT(\bar G_\gamma)\,[\mathbb{B}\circ \bar G_\gamma] \,, \tag{2.3}$$

in distribution, where α_k denotes the uniform empirical process and $\mathbb{B}$ a Brownian bridge, and

$$\lim_{n\to+\infty} \frac{T(\bar G_{\gamma+a_n}) - T(\bar G_\gamma)}{a_n} = DT(\bar G_\gamma)\left[\frac{\partial \bar G_\gamma}{\partial\gamma}\right]. \tag{2.4}$$

Clearly, condition (A$_2$) requires that $D_{\gamma,\rho}$, $\mathbb{B}\circ\bar G_\gamma$ and $\frac{\partial \bar G_\gamma}{\partial\gamma}$ are in $L^1(\mathbb{R}^+, \mu_{1,\gamma}) \cap L^1(\mathbb{R}^+, \mu_{2,\gamma})$.

(**A**$_3$) A scale invariance property, namely for all $\bar F$ such that $T(\bar F)$ is defined and all $\sigma > 0$,

$$T_1\left(\bar F\left(\frac{\bullet}{\sigma}\right)\right) = T_1(\bar F) \qquad \text{and} \qquad T_2\left(\bar F\left(\frac{\bullet}{\sigma}\right)\right) = \sigma\, T_2(\bar F) \,. \tag{2.5}$$

As in the Introduction, we use the notation $\sigma_n = \sigma(u_n)$. Then, if we denote by $(\gamma^b_n, \sigma^b_n) = T(\bar F_{u_n})$ the values of γ and σ obtained when $\bar G_{\gamma,\sigma_n}$ has been substituted by the true distribution function $\bar F_{u_n}$ of the excesses, we deduce from (2.2)–(2.5) that

$$\frac{1}{a_n}\left(\gamma^b_n - \gamma,\, \frac{\sigma^b_n}{\sigma_n} - 1\right) \longrightarrow DT(\bar G_\gamma)[D_{\gamma,\rho}] =: \big(\mathcal{L}_1(\gamma,\rho),\, \mathcal{L}_2(\gamma,\rho)\big) \,, \tag{2.6}$$

and

$$\sqrt{N_n}\left(\hat\gamma_n - \gamma^b_n,\, \frac{\hat\sigma_n}{\sigma_n} - \frac{\sigma^b_n}{\sigma_n}\right) \xrightarrow{d} \tag{2.7}$$

$$\xrightarrow{d} \left(\int_0^\infty \mathbb{B}\circ\bar G_\gamma \; d\mu_{1,\gamma} \,,\, \int_0^\infty \mathbb{B}\circ\bar G_\gamma \; d\mu_{2,\gamma}\right) =: (Z_1, Z_2) \,.$$

Consequently,

$$\sqrt{N_n}\left(\hat\gamma_n - \gamma,\, \frac{\hat\sigma_n}{\sigma_n} - 1\right) \xrightarrow{d} (Z_1, Z_2) + \lambda\big(\mathcal{L}_1(\gamma,\rho),\, \mathcal{L}_2(\gamma,\rho)\big) \,. \tag{2.8}$$

Note that (2.6) contains the bias of approximation of γ and σ, whereas (2.7) involves the limiting distribution of $\hat\gamma_n$ and $\hat\sigma_n$. This shows that the bias of approximation on the parameters is of order a_n and under (1.5), $\hat\gamma_n$ and $\hat\sigma_n/\sigma_n$ are asymptotically biased when $\lambda \neq 0$. In this paper, we will focus on the *asymptotic functional bias of approximation*, defined as:

$$A_E(x) := \lim_{n\to\infty} \frac{\bar{F}^*_{u_n}(x) - \bar{G}_{\gamma_n^b,\,\sigma_n^b/\sigma_n}(x)}{a_n} .$$

This quantity is important since it measures the first order non stochastic discrepancy between the unknown target tail function, $\bar{F}^*_{u_n}$, and its observable counterpart, $\bar{G}_{\hat{\gamma}_n,\hat{\sigma}_n/\sigma_n}$. This bias is important to statisticians who are more interested in estimating small tail probabilities than in estimating γ (as Drees (1998) and Drees *et al.* (2004) who have studied the asymptotic behaviour of the maximum likelihood estimators $(\hat{\gamma}_n, \hat{\sigma}_n)$).

Using (1.4) and a Taylor expansion, it can easily be proved that

$$A_E(x) \;=\; D_{\gamma,\rho}(x) \;-\; \mathcal{L}_1(\gamma,\rho)\,\frac{\partial \bar{G}_\gamma}{\partial\gamma}(x) \;+\; \mathcal{L}_2(\gamma,\rho)\,x\,\frac{\partial \bar{G}_\gamma}{\partial x}(x) \;. \tag{2.9}$$

This expression contains both the bias of approximation (1.4) and the error of approximation on the parameters (2.6).

This result has been first established in Diebolt *et al.* (2003) in the special case of the probability-weighted moments estimators of Hosking and Wallis (1987).

Our main result is summarized in the following theorem.

Theorem 1. *Under assumptions* (A_1)–(A_3) *and (1.5), we have*

$$A_E \equiv 0 \qquad \text{when} \;\; \rho = 0 \;.$$

Proof: From (A_1), we clearly have that $\frac{\partial T(\bar{G}_{\gamma,\sigma})}{\partial\gamma} = (1,0)$ and from (2.4) we deduce that $\frac{\partial T(\bar{G}_{\gamma,1})}{\partial\gamma} = DT(\bar{G}_\gamma)\Big[\frac{\partial \bar{G}_\gamma}{\partial\gamma}\Big]$. Therefore, in the case $\rho = 0$, since $D_{\gamma,0} = \frac{\partial \bar{G}_\gamma}{\partial\gamma}$, we derive that $\mathcal{L}_1(\gamma,0) = 1$ and $\mathcal{L}_2(\gamma,0) = 0$. Consequently, $A_E(x) = D_{\gamma,0}(x) - \frac{\partial \bar{G}_\gamma}{\partial\gamma}(x) = 0$ in the case $\rho = 0$. This explains why as soon as we use estimators which satisfy (A_1)–(A_3), the function A_E becomes the null function for $\rho = 0$. □

This result is particularly remarkable. Indeed:

- This result means that in the case $\rho = 0$, although the bias of $\widehat{\gamma}_n$ is of order a_n, its contribution is cancelled by compensation with $D_{\gamma,0}$ in the expression of the function A_E. It can be seen that A_E remains small whenever $|\rho|$ is small.

- The second order parameter ρ is zero for many usual distributions in the Gumbel domain of attraction ($\gamma = 0$): e.g., the normal, lognormal, gamma and classical Weibull distributions. Hence, our result applies to all of these distributions. In the Frechet domain of attraction, we also have distributions with $\rho = 0$, such as the loggamma distribution.
- This result is closely linked to penultimate approximation in Worms (2002).
- This remarkable behaviour of the error of functional approximation A_E when $\rho = 0$ contrasts strongly with theoretical and experimental results concerning the semiparametric estimators of γ, for example the Hill estimator (Hill, 1975). Indeed, the bias of these estimators explodes when ρ tends to 0. This motivates many recent works on exponential regression model, where estimators with reduced bias are proposed (see, for example, Beirlant *et al.*, 1999).

3. MAXIMUM LIKELIHOOD ESTIMATION FOR $\gamma > 0$

In this section, we will prove the following theorem.

Theorem 2. *For $\gamma > 0$, the maximum likelihood estimators are regular in the sense that they satisfy conditions (A_1)–(A_3).*

With this aim, we first establish the local existence and unicity of these estimators. Then, we prove the regularity of T in the maximum likelihood setting with $\gamma > 0$.

We will need some notation. First, we denote by $g_\theta = g_{\gamma,\sigma}$ the density of the GPD distribution with parameters θ. Then the score function $s_\theta(x)$ is the gradient of $\ln g_\theta(x)$ with respect to θ. It is a function of x taking its values in $\mathbb{R}^2$. The score function $s_\theta(x)$ is of the form

$$s_\theta(x) \;=\; \big(s_1(x;\theta),\, s_2(x;\theta)\big)^T$$

where, denoting $y = x/\sigma$,

$$s_1(x;\gamma,\sigma) \;=\; \begin{cases} \dfrac{(1+\gamma y)\,\ln(1+\gamma y) - (\gamma+1)\gamma y}{\gamma^2(1+\gamma y)} & \text{if } \;\gamma > 0\,, \\[2ex] \dfrac{y(y-2)}{2} & \text{if } \;\gamma = 0\,, \end{cases} \tag{3.1}$$

and

$$s_2(x;\gamma,\sigma) \;=\; \frac{y-1}{\sigma(1+\gamma y)} \qquad \text{if } \;\gamma \geq 0\,. \tag{3.2}$$

Let also $\psi_\theta(x)$ denote the derivative of $s_\theta(x)$ with respect to x:

$$\psi_\theta(x) := (s_\theta)'(x) := \big(\psi_1(x;\theta),\, \psi_2(x;\theta)\big)^T = \left(\frac{x-\sigma}{(\sigma+\gamma x)^2},\, \frac{1+\gamma}{(\sigma+\gamma x)^2}\right) .$$

3.1. Existence and local unicity

We consider the Küllback–Leibler divergence between two densities of probabilities h and g related to a measure of reference ν:

$$\mathrm{Ent}_g\left(\frac{h}{g}\right) := \int \ln\left(\frac{h(x)}{g(x)}\right) h(x)\; \nu(dx) .$$

It takes, by convention, the value ∞ when the integral is not finite. Let $\theta \in \Theta = \{(\gamma,\sigma)\colon \gamma > 0,\ \sigma > 0\}$. Under our assumptions, F admits a density f and the two Küllback–Leibler divergences between g_θ and f are defined as (note that they can take the value ∞)

$$d_{KL}(g_\theta \,|\, f) := \mathrm{Ent}_f\left(\frac{g_\theta}{f}\right) = \int_0^\infty \ln\left(\frac{g_\theta(x)}{f(x)}\right) g_\theta(x)\; dx$$

and

$$d_{KL}(f \,|\, g_\theta) := \int_0^\infty \ln\left(\frac{f(x)}{g_\theta(x)}\right) f(x)\; dx .$$

These quantities are ≥ 0 and $d_{KL}(f|\, g) = 0$ if and only if $f = g$ a.e. Similar properties exist for $d_{KL}(g|\, f)$. Splitting the logarithm into two parts, we obtain

$$d_{KL}(f \,|\, g_\theta) = \int_0^\infty \ln f(x)\, f(x)\; dx \;+\; \Delta(\theta, \bar{F}) ,$$

where

$$\Delta(\theta,\, \bar{F}) := \int_0^\infty \ln g_\theta(x)\; d\bar{F}(x) = -\int_0^\infty \ln g_\theta(x)\, f(x)\; dx . \tag{3.3}$$

Thus, $d_{KL}(f|\, g_\theta)$ is minimal if and only if $\Delta(\theta, \bar{F})$ is minimal. Here, the family $\{g_\theta\colon \theta \in \Theta\}$ is identifiable: the distance d_{KL} between g_θ and $g_{\theta'}$ is equal to zero if and only if $\theta = \theta'$.

For each $\bar{F}$, the function $\theta \mapsto \Delta(\theta, \bar{F})$ is continuous on Θ as soon as

$$\int_1^\infty \ln x\;\, f(x)\; dx < \infty . \tag{C$_1$}$$

Lemma 1 below guarantees the local existence of the maximum likelihood estimators.

Lemma 1. *Under (C_1), the restriction of $\theta \mapsto \Delta(\theta, \bar{F})$ to each closed ball $\mathcal{K}$ contained in Θ reaches its minimum value in $\mathcal{K}$ and at each point where this minimum is reached, we have*

$$\int_0^\infty s_\theta(x)\, d\bar{F}(x) \;=\; \underline{0}\ . \tag{3.4}$$

The proof of this lemma is straightforward. Remark that (3.4) constitutes the likelihood equations.

We now consider the local unicity. First define

$$W(\theta, \bar{F}) \;:=\; \int_0^\infty \bar{F}(x)\, \psi_\theta(x)\, dx\ .$$

Integrating by parts yields the following result:

Lemma 2. *Under (C_1),*

$$\bar{F}(x)\, \ln x \underset{x\to\infty}{\longrightarrow} 0\ , \tag{C$_2$}$$

and if $\bar{F}(0) = 1$, we have

$$W(\theta, \bar{F}) \;=\; \left(0, \frac{1}{\sigma}\right)^T - \int_0^\infty s_\theta(x)\, d\bar{F}(x)\ . \tag{3.5}$$

Remark that if (C_2) is satisfied, then (C_1) can be rewritten as $\int_1^\infty (\bar{F}(x)/x)\, dx < \infty$. Moreover, if we assume the mild condition that there exists an $\varepsilon > 0$ such that

$$\bar{F}(x)\, (\ln x)^{1+\varepsilon} \underset{x\to\infty}{\longrightarrow} 0\ , \tag{C$_3$}$$

then (C_1) and (C_2) are satisfied. Note that in Appendix (4.3), we will prove that (C_3) is satisfied by $\bar{F}^*_{u_n}$.

In all the sequel, we use the notation $\theta^\star = (\gamma^\star, 1)$, $\gamma^\star > 0$ denoting the true value of γ and $\bar{G} := \bar{G}_{\theta^\star}$. The local unicity is established in the following lemma.

Lemma 3. *There exists a closed ball $\mathcal{V}^\star$ centered at $\theta^\star$ such that the restriction of the application $\theta \mapsto \Delta(\theta, \bar{F}^*_{u_n})$ to $\mathcal{V}^\star$ is strictly convex for all n sufficiently large, as $\bar{F}^*_{u_n}$ converges to $\bar{G}$ in the sense described in Worms (2000, 2002).*

Proof of Lemma 3: We consider the second order differential $D^2 \ln g_\theta(x)$ with respect to θ, represented by a $2{\times}2$ matrix-valued function of x. For each suitable $\bar{F}$ and each $\theta \in \Theta$, the matrix

$$I(\theta, \bar{F}) := \int_0^\infty D^2 \ln g_\theta(x)\, d\bar{F}(x)$$

is a Fisher-type information matrix. Recall that the symmetric matrix $I(\theta, \bar{G}_\theta)$ is definite positive for each $\gamma > -1/2$.

We show in Appendix 4.1 via an integration by parts that the matrix $I(\theta, \bar{F}^*_{u_n})$ converges, uniformly in θ in some compact neighbourhood $\mathcal{V}^\star$ of $\theta^\star$, to a matrix $I(\theta, \bar{G})$, as $n \to \infty$. Consequently, $I(\theta, \bar{F}^*_{u_n})$ is definite positive for all $\theta \in \mathcal{V}^\star$ and all n sufficiently large. In this case, we have for all n sufficiently large, unicity of the minimum of $\Delta(\theta, \bar{F}^*_{u_n})$ for $\theta \in \mathcal{V}^\star$, i.e. local unicity. We denote by $\theta_n = T(\bar{F}^*_{u_n})$ the point in $\mathcal{V}^\star$ minimizing $\Delta(\theta, \bar{F}^*_{u_n})$. □

Remark that the functional T is sequentially continuous in the following sense. Since the $T(\bar{F}^*_{u_n})$'s are in the compact $\mathcal{V}^\star$, the sequence that they form admits at least an adherence value, which belongs to $\mathcal{V}^\star$. We deduce from compactness and identifiability that any adherence value of this sequence is necessarily $\theta^\star$. It therefore follows that $T(\bar{F}^*_{u_n})$ converges to $T(\bar{G}) = \theta^\star$.

3.2. Regularity of T in the maximum likelihood case

In this section, we will essentially prove the Hadamard differentiability of T, with a differential given by

$$H \longmapsto \Big[I\big(T(\bar{G}), \bar{G}\big)\Big]^{-1} \int_0^\infty H(x)\, \psi_{T(\bar{G})}(x)\, dx \; . \tag{3.6}$$

Starting from (3.5), we obtain

$$\int_0^\infty \bar{F}^*_{u_n}(x)\, D\psi_\theta(x)\, dx = \begin{bmatrix} 0 & 0 \\ 0 & -\frac{1}{\sigma^2} \end{bmatrix} - \int_0^\infty Ds_\theta(x)\, d\bar{F}^*_{u_n}(x)$$

$$=: M - \int_0^\infty Ds_\theta(x)\, d\bar{F}^*_{u_n}(x) \; ,$$

that is

$$I(\theta, \bar{F}^*_{u_n}) = -\int_0^\infty \bar{F}^*_{u_n}(x)\, D\psi_\theta(x)\, dx + M \; .$$

By local existence and unicity, for all n sufficiently large

$$-\left(0, \frac{1}{\sigma}\right)^T + W(\theta, \bar{F}^*_{u_n}) = \underline{0} \qquad \text{if and only if} \quad \theta = T(\bar{F}^*_{u_n}) \; .$$

Therefore

$$-\Big(0,\ \frac{1}{T_2(\bar F^*_{u_n})}\Big)^T + W\big(T(\bar F^*_{u_n}), \bar F^*_{u_n}\big) \;=\; \underline{0}$$

and

$$-\Big(0,1\Big)^T + W\big(T(\bar G), \bar G\big) \;=\; \underline{0}\ .$$

Thus,

$$\begin{aligned}
\underline{0} \;&=\; \frac{1}{a_n}\Bigg[W\big(T(\bar F^*_{u_n}), \bar F^*_{u_n}\big) \;-\; W\big(T(\bar G), \bar G\big) \;-\; \Big(0,\ \frac{1}{T_2(\bar F^*_{u_n})}\Big)^T + \Big(0,1\Big)^T\Bigg] \\
&=\; \frac{1}{a_n}\Bigg[\int_0^\infty \bar F^*_{u_n}(y)\,\psi_{T(\bar F^*_{u_n})}(y)\;dy \;-\; \int_0^\infty \bar G(y)\,\psi_{T(\bar G)}(y)\;dy \\
&\qquad -\; \Big(0,\ \frac{1}{T_2(\bar F^*_{u_n})}\Big)^T + \Big(0,1\Big)^T\Bigg]\ . \qquad (3.7)
\end{aligned}$$

We know that $\lim_{n\to\infty} T(\bar F^*_{u_n}) \;=\; T(\bar G) \;=\; (\gamma^\star, 1)$. For each y, we use a Taylor expansion of order 2 with integral remainder of $\psi_\theta(y)$ around $\psi_{T(\bar G)}(y)$:

$$\psi_{T(\bar F^*_{u_n})}(y) \;=\; \psi_{T(\bar G)}(y) + D\psi_\theta(y)\big|_{\theta=(\gamma^\star,1)}\big[T(\bar F^*_{u_n}) - T(\bar G)\big] \;+\; \text{remainder}\,. \qquad (3.8)$$

Recall the principle of Taylor expansions of order 2 with integral remainder. Let f be a function of two variables. Denoting $g(t) = f(a_1 + th_1,\ a_2 + th_2)$ and using

$$g(1) \;=\; g(0) \;+\; g'(0) \;+ \int_0^1 (1-t)\, g''(t)\;dt\ ,$$

it follows that

$$\begin{aligned}
f\big(a_1 + h_1,\ a_2 + h_2\big) \;&=\; \\
&= f(a_1, a_2) \;+\; h_1\,\partial_1 f(a_1,a_2) \;+\; h_2\,\partial_2 f(a_1,a_2) \\
&\quad + \int_0^1 (1-t)\,\big(h_1^2\,\partial_{11} f + 2\,h_1\,h_2\,\partial_{12} f + h_2^2\,\partial_{22} f\big)\,(a_1 + th_1,\ a_2 + th_2)\;dt\ .
\end{aligned}$$

We will apply this formula to the two functions $\psi_j(x;\gamma,\sigma)$, $j=1,2$, at a fixed x, as functions of (γ,σ). The point (a_1,a_2) will be $(\gamma^\star,1)$ and the point $(a_1{+}h_1, a_2{+}h_2)$ will be (γ,σ) close to $(\gamma^\star,1)$. Denote by $\Delta\gamma := \gamma - \gamma^\star$ and $\Delta\sigma := \sigma - 1$. We have ($\partial_1$ denoting the derivative in γ, and ∂_2 the derivative in σ)

$$\partial_1\psi_1(x;\,\gamma,\sigma) \;=\; -\,\frac{2x\,(x-\sigma)}{(\sigma+\gamma x)^3} \qquad\text{and}\qquad \partial_2\psi_1(x;\,\gamma,\sigma) \;=\; -\,\frac{(\gamma+2)x-\sigma}{(\sigma+\gamma x)^3}\ ,$$

$$\partial_1\psi_2(x;\,\gamma,\sigma) \;=\; -\,\frac{(\gamma+2)x-\sigma}{(\sigma+\gamma x)^3} \qquad\text{and}\qquad \partial_2\psi_2(x;\,\gamma,\sigma) \;=\; -\,\frac{2(1+\gamma)}{(\sigma+\gamma x)^3}\ .$$

We have similar expressions for the second order partial derivatives. As x tends to infinity, there are all of order $\mathcal{O}(1)$. For example,

$$\partial_{11}\psi_1(x;\,\gamma,\sigma) \;=\; \frac{6x^2(x-\sigma)}{(\sigma+\gamma x)^4}\;.$$

Thus

$$(\Delta\gamma)^2\int_0^1 (1-t)\,\partial_{11}\psi_1\big(x;\,\gamma^\star+t\,\Delta\gamma,\,1+t\,\Delta\sigma\big)\;dt \;=$$
$$=\;(\Delta\gamma)^2\int_0^1 (1-t)\,\frac{6x^2(x-1-t\,\Delta\sigma)}{\big[\,1+\gamma^\star x+(\Delta\sigma+x\,\Delta\gamma)\,t\,\big]^4}\;\;dt\;.$$

We are interested in

$$\int_0^\infty \bar F^*_{u_n}(y)\,\psi_{T(\bar F^*_{u_n})}(y)\;dy \;-\int_0^\infty \bar G(y)\,\psi_{T(\bar G)}(y)\;dy\;.$$

In the development $\psi_{T(\bar F^*_{u_n})}(y)-\psi_{T(\bar G)}(y)$, the contribution to the integral remainder due to $\partial_{11}\psi_1$ is of the form

$$(\Delta\gamma)^2\int_0^\infty \bar F^*_{u_n}(y)\,\int_0^1 (1-t)\,\frac{6y^2(y-1-t\,\Delta\sigma)}{\big[\,1+\gamma^\star y+(\Delta\sigma+y\,\Delta\gamma)\,t\,\big]^4}\;\;dt\,dy\;,$$

which is

$$(\Delta\gamma)^2\int_0^1 (1-t)\,\int_0^\infty \bar F^*_{u_n}(y)\,\frac{6y^2(y-1-t\,\Delta\sigma)}{\big[\,1+\gamma^\star y+(\Delta\sigma+y\,\Delta\gamma)\,t\,\big]^4}\;\;dy\,dt\;.$$

We now show that this term is $O((\Delta\gamma)^2)$. The range y small is trivial, since the dominating term in the ratio is $-6y^2$. Therefore, we will look only at the range $y\ge y_0>0$. We have therefore to study

$$\int_0^1 (1-t)\,\int_{y_0}^\infty \bar F^*_{u_n}(y)\,\frac{6\big(1-(1+t\,\Delta\sigma)/y\big)}{y\big[\,1/y+\gamma^\star+(\Delta\sigma/y+\Delta\gamma)\,t\,\big]^4}\;\;dy\,dt\;,$$

which can be reduced to study the quantity

$$\int_0^1 (1-t)\,\int_{y_0}^\infty \frac{\bar F^*_{u_n}(y)}{y}\;\;dy\,dt\;,$$

which is integrable under our conditions and converges to $\int_0^1(1-t)\int_{y_0}^\infty(\bar G(y)/y)\,dy\,dt$. It is thus a bounded sequence.

From (3.7) and (3.8), we obtain that

$$\left(I(\theta^\star, \bar F^*_{u_n}) - \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{T_2(\bar F^*_{u_n})} - 1 \end{bmatrix} \right) \left[\frac{T(\bar F^*_{u_n}) - T(\bar G)}{a_n} \right] =$$

$$= \int_0^\infty \frac{\bar F^*_{u_n}(y) - \bar G(y)}{a_n}\, \psi_{\theta^\star}(y)\, dy \;+\; \text{remainder} .$$

Using the Appendix 4.3, we have

$$\int_0^\infty \frac{\bar F^*_{u_n}(y) - \bar G(y)}{a_n}\, \psi_{\theta^\star}(y)\, dy \;\underset{n\to\infty}{\longrightarrow}\; \int_0^\infty D_{\gamma^\star,\rho}(y)\, \psi_{\theta^\star}(y)\, dy .$$

Moreover, $T_2(\bar F^*_{u_n}) \to 1$ and we have established that $I(\theta^\star, \bar F^*_{u_n}) \to I(\theta^\star, \bar G)$ which is definite positive, thus

$$I(\theta^\star, \bar F^*_{u_n}) - \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{T_2(\bar F^*_{u_n})} - 1 \end{bmatrix}$$

is inversible for all n sufficiently large. We conclude therefore that

$$\begin{aligned} \frac{T(\bar F^*_{u_n}) - T(\bar G)}{a_n} &= \left[I(\theta^\star, \bar F^*_{u_n}) - \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{T_2(\bar F^*_{u_n})} - 1 \end{bmatrix} \right]^{-1} \\ &\quad \cdot \int_0^\infty \frac{\bar F^*_{u_n}(y) - \bar G(y)}{a_n}\, \psi_{\theta^\star}(y)\, dy \;+\; \text{remainder} \quad \longrightarrow \\ &\quad \longrightarrow \; \Big[I\big(T(\bar G), \bar G\big) \Big]^{-1} \int_0^\infty D_{\gamma^\star,\rho}(y)\, \psi_{\theta^\star}(y)\, dy . \end{aligned} \tag{3.9}$$

Therefore (3.6) is now established for $\bar F^*_{u_n}$ and a similar study is carried out for sequences $\big(\bar F^*_{u_n} + k_n^{-1/2} \alpha_{k_n}(\bar F^*_{u_n})\big)_{n\ge1}$ in Appendices 4.4 and 4.5. □

We have now to compute the couple of biases, to check that the principal term of the functional bias of approximation cancels for $\rho = 0$, and to compute it for $\rho \neq 0$. We know that

$$I^{-1} := \Big[I\big(T(\bar G), \bar G\big) \Big]^{-1} = (\gamma + 1) \begin{bmatrix} \gamma + 1 & -1 \\ -1 & 2 \end{bmatrix} .$$

The column 2×1 of biases (first the bias concerning γ, then the bias concerning σ close to 1, i.e. for σ/σ_n), is

$$\frac{a_n}{\big(1+|\rho|\big)\,\big(1+\gamma+|\rho|\big)} \begin{bmatrix}\gamma+1\\ |\rho|\end{bmatrix} . \tag{3.10}$$

This formula is also applicable, at least formally, in the case $\gamma = 0$. When we choose $\rho = 0$, we find $a_n\,(1,0)^T$, as expected.

Remark 1. If we consider the Hall model (1982) defined by

$$1-F(t) \;=\; C\,t^{-1/\gamma}\Big(1+D\,t^{\rho/\gamma}\big(1+o(1)\big)\Big) \;, \qquad t\to\infty \;,$$

where $\rho<0$, $C>0$ and $D\in\mathbb{R}$, direct computations lead to $A(t) = \rho\,D\,C^{\rho}(\gamma+\rho)\cdot t^{\rho}(1+o(1))$ and so $a_n = \rho\,D\,(\gamma+\rho)\,u_n^{\rho/\gamma}(1+o(1))$. The vector of biases (3.10) is the one given in Smith (1987). Remark that this verification is direct for the bias of γ. On the other hand, it is not the case for the bias of $\sigma\big/\sigma_n$, taking into account the fact that Smith (1987) took a σ_n different from ours. It is thus necessary to take this difference into account. In the same way, Drees *et al.* (2004) have also, but in a different way, obtained the vector of biases for a standardization different from ours and from that of Smith (1987), but as previously mentioned a repercussion of this difference gives again (3.10).

Remark 2. We have just shown that the maximum likelihood estimators satisfy conditions (A$_1$)–(A$_3$) with $\mu_{k,\gamma}(dx) = \psi_k(x;\gamma,1)\,dx$, $k=1,2$.

Since $(\gamma_n^b,\sigma_n^b) = T(\bar F_{u_n})$, we have $(\gamma_n^b,\frac{\sigma_n^b}{\sigma_n}) = T(\bar F^*_{u_n})$. Then, by (3.6) and (3.7),

$$\begin{aligned}
T(\bar F^*_{u_n}) - T(\bar G_{\gamma,1}) \;&=\; \left(\gamma_n^b-\gamma,\, \frac{\sigma_n^b}{\sigma_n}-1\right)\\
&=\; a_n\,I^{-1}\int_0^{\infty} D_{\gamma,\rho}(x)\,\psi_{\gamma,1}(x)\;dx \;+\; o(a_n)\\
&=\; \frac{a_n}{\big(1+|\rho|\big)\,\big(1+\gamma+|\rho|\big)}\begin{bmatrix}\gamma+1\\ |\rho|\end{bmatrix} \;+\; o(a_n)\;.
\end{aligned}$$

It follows that

$$\begin{aligned}
&\bar F^*_{u_n}(x) - \bar G_{\gamma_n^b}\left(\frac{\sigma_n^b\,x}{\sigma_n}\right) \;=\\
&\quad=\; a_n\left(D_{\gamma,\rho}(x) + \frac{\gamma+1}{(1{+}|\rho|)\,(1{+}\gamma{+}|\rho|)}\,\frac{\partial G_\gamma}{\partial\gamma}(x) - \frac{|\rho|\,x}{(1{+}|\rho|)\,(1{+}\gamma{+}|\rho|)}\,\frac{\partial G_\gamma}{\partial x}(x)\right)\\
&\qquad + o(a_n)\;.
\end{aligned}$$

Therefore, it can be directly checked that, when $\rho{=}0$, $A_E(x){=}0$ for all x.

4. APPENDIX

4.1. Convergence of the information matrices for $\bar{F}^*_{u_n}$

We use the following notations in all the sequel: $\bar{G} := \bar{G}_{\theta^\star}$ and $\Psi(\cdot\,|\theta) := D\psi_\theta(x)$. We would like to prove that, $\forall\,\varepsilon > 0$,

$$\left| \int_0^\infty \bar{G}(y)\,\Psi(y|\theta)\;dy - \int_0^\infty \bar{F}^*_{u_n}(y)\,\Psi(y|\theta)\;dy \right| \;\leq\; \varepsilon\;,$$

for n sufficiently large, $\forall\,\theta \in \mathcal{V}^\star = \mathcal{V}(\theta^\star)$.

This proof can be divided into two parts. First, we consider an $\varepsilon > 0$ and we will prove that it is possible to choose a number A such that each quantity

$$\left| \int_A^\infty \bar{G}(y)\,\Psi(y|\theta)\;dy \right| \qquad \text{and} \qquad \left| \int_A^\infty \bar{F}^*_{u_n}(y)\,\Psi(y|\theta)\;dy \right|$$

is smaller than $\varepsilon/2$, $\forall\,\theta \in \mathcal{V}^\star$ and $\forall\,n \geq n_1(A)$.

Then, using this number A, we will establish our result by reasoning on

$$\int_0^A \bar{F}^*_{u_n}(y)\,\Psi(y|\theta)\;dy$$

and using the mean value theorem.

In order to establish the first part of this result, we use the classical change of variables

$$\left| \int_A^\infty \bar{G}(y)\,\Psi(y|\theta)\;dy \right| \;=\; \left| \int_B^\infty e^{-s}\,\Psi\big(g(s)\big|\theta\big)\,g'(s)\;ds \right|$$

and

$$\left| \int_A^\infty \bar{F}^*_{u_n}(y)\,\Psi(y|\theta)\;dy \right| \;=\; \left| \int_{B_n}^\infty e^{-s}\,\Psi\big(g_n(s)\big|\theta\big)\,g_n'(s)\;ds \right|\;, \tag{4.1}$$

where $g_n(s) = (e^{\gamma^\star s} - 1)/\gamma^\star + a_n R_n(s)$ and $g(s) = (e^{\gamma^\star s} - 1)/\gamma^\star$, and we prove that the latter quantity is smaller than $\varepsilon/2$ for an $\varepsilon = \varepsilon(B_n)$, uniformly in $\theta = (\gamma,\sigma)$ close to $\theta^\star = (\gamma^\star, 1)$, and for n sufficiently large. Remark that the first quantity can be treated similarly and recall that $g_n'(s)$ is of the form $e^{\gamma^\star s} + a_n r_n(s)$, an expression that we will use later on. All the proof will be done in the case $\gamma^\star > 0$ and we will use the first component of $\Psi(\cdot|\theta)$, the other ones can be treated similarly (and in fact more easily). Thus, we have to bound the quantity

$$\frac{g_n(s)\,|g_n(s) - \sigma|}{\big(\sigma + \gamma g_n(s)\big)^3}\;|g_n'(s)| \qquad \text{for} \quad s \geq B_n\;.$$

Recall that $y = \bar{F}_{u_n}^{*\,-1}(\mathrm{e}^{-s})$. Since $\bar{F}_{u_n}^{*\,-1}$ and $s \mapsto \mathrm{e}^{-s}$ are decreasing, y is increasing. Consequently, since $g_n(s) \nearrow \infty$ for $s \nearrow \infty$ $\forall\, n$, we can use the bound

$$\mathrm{cst}\, \frac{g_n^2(s)}{g_n^3(s)}\, |g_n'(s)|$$

uniformly for $\theta \in \mathcal{V}^\star$. We have therefore to study

$$\int_{B_n}^{\infty} e^{-s}\, \frac{|g_n'(s)|}{g_n(s)}\, ds \;=\; \int_{B_n}^{\infty} e^{-s}\, \frac{g_n'(s)}{g_n(s)}\, ds \tag{4.2}$$

since $g_n' \ge 0$. By integrating by parts, this integral is equal to

$$-\, e^{-B_n} \ln\bigl(g_n(B_n)\bigr) \;+ \int_{B_n}^{\infty} e^{-s} \ln\bigl(g_n(s)\bigr)\; ds\;. \tag{4.3}$$

Using the fact that the first term is negative and that if $s \ge B_n$ then $g_n(s) \ge g_n(B_n)$, we obtain the first part of the proof.

Concerning the second part of the proof, we first prove that

$$\int_0^A \bar{F}_n^*(y)\, \Psi(y|\theta)\; dy \;\longrightarrow\; \int_0^A \bar{G}(y)\, \Psi(y|\theta)\; dy$$

as $n \to \infty$, uniformly in θ close to $\theta^\star = (\gamma^\star, 1)$. We have

$$\int_0^A \bar{F}_n^*(y)\, \Psi(y|\theta)\; dy \;=\; \int_0^{B_n} e^{-s}\, \Psi\bigl(g_n(s)\big|\theta\bigr)\, g_n'(s)\; ds$$

and

$$\int_0^A \bar{G}(y)\, \Psi(y|\theta)\; dy \;=\; \int_0^{B} e^{-(1-\gamma^\star)s}\, \Psi\bigl(g(s)\big|\theta\bigr)\; ds\;.$$

Thus we look at

$$\begin{aligned}
\sup_{\theta\in\mathcal{V}^\star} \left| \int_0^A \bar{F}_n^*(y)\, \Psi(y|\theta)\; dy - \int_0^A \bar{G}(y)\, \Psi(y|\theta)\; dy \right| \;=&\\
= \sup_{\theta\in\mathcal{V}^\star} \Bigg| \int_0^B e^{-(1-\gamma^\star)s} \Big[\Psi\bigl(g_n(s)\big|\theta\bigr) - \Psi\bigl(g(s)\big|\theta\bigr)\Big]\, ds&\\
+\, a_n \int_0^B e^{-s}\, r_n(s)\, \Psi\bigl(g_n(s)\big|\theta\bigr)\; ds \Bigg|&\\
+ \sup_{\theta\in\mathcal{V}^\star} \left| \int_B^{B_n} e^{-s}\, \Psi\bigl(g_n(s)\big|\theta\bigr)\, g_n'(s)\; ds \right|&\\
= \sup_{\theta\in\mathcal{V}^\star} \Bigg| \int_0^B e^{-(1-\gamma^\star)s}\, a_n\, R_n(s)\, \frac{\partial\Psi}{\partial x} \left(\frac{e^{\gamma^\star s} - 1}{\gamma^\star} + \xi_n(s) \,\middle|\, \theta \right) ds&\\
+\, a_n \int_0^B e^{-s}\, r_n(s)\, \Psi\bigl(g_n(s)\big|\theta\bigr)\; ds \Bigg|&\\
+ \sup_{\theta\in\mathcal{V}^\star} \left| \int_B^{B_n} e^{-s}\, \Psi\bigl(g_n(s)\big|\theta\bigr)\, g_n'(s) \right|\,,&
\end{aligned}$$

where $\xi_n(s) \in (0, a_n R_n(s))$ and where $\mathcal{V}^\star$ is a compact neighbourhood of $\theta^\star$. Therefore, we have to study two supremum separately. We have some very useful properties (Potter bounds) that we recall below (see Worms, 2000):

- $$\begin{aligned} &\frac{V\big(s + V^{-1}(u_n)\big) - u_n}{\sigma_n} \;=\; \frac{e^{\gamma^\star s} - 1}{\gamma^\star} + a_n R_n(s) \;, \qquad a_n \in \mathbb{R} \;, \\ &|R_n(s)| \;\le\; \text{cst}\; e^{(\gamma^\star + \eta)s} \int_0^s e^{\rho t}\, dt \;, \\ &R_n(s) \underset{n\to\infty}{\longrightarrow} \int_0^s e^{\gamma^\star z} \int_0^z e^{\rho t}\, dt\, dz \;. \end{aligned}$$

- $$\begin{aligned} &\frac{V'\big(s + V^{-1}(u_n)\big)}{\sigma_n} \;=\; e^{\gamma^\star s} + a_n r_n(s) \;, \\ &|r_n(s)| \;\le\; \text{cst}\; e^{(\gamma^\star + \eta)s} \int_0^s e^{\rho t}\, dt \;, \\ &r_n(s) \underset{n\to\infty}{\longrightarrow} e^{\gamma^\star s} \int_0^s e^{\rho t}\, dt \;. \end{aligned}$$

For the first supremum, we study the two quantities of the sum separately. Concerning the first one and taking into account the fact that

$$\sup_{\theta \in \mathcal{V}^\star,\, x \in [0,A]} \big|\Psi(x|\theta)\big| \;<\; \infty \qquad \text{and} \qquad \sup_{\theta \in \mathcal{V}^\star,\, x \in [0,A]} \left|\frac{\partial \Psi}{\partial x}(x|\theta)\right| \;<\; \infty \;,$$

we can restrict ourself to a compact in x, i.e. work with the product of the two compacts. This is what we do below, where we have to study

$$\begin{aligned} &\left| e^{-(1-\gamma^\star)s}\, a_n\, R_n(s)\, \frac{\partial \Psi}{\partial x} \left(\frac{e^{\gamma^\star s} - 1}{\gamma^\star} + \xi_n(s) \,\middle|\, \theta \right) \right| \;\le \\ &\qquad\qquad \le\; \text{cst}\; |a_n|\, e^{-(1 - 2\gamma^\star - \eta)s} \sup_{\theta \in \mathcal{V}^\star,\, x \in [0,A]} \left| \frac{\partial \Psi}{\partial x}(x|\theta) \right| \;, \end{aligned}$$

taking the fact that $|R_n(s)| \le \mathrm{K}\, e^{(\gamma^\star + \eta)s}$ into account. Consequently, the first quantity in the first supremum tends to 0 uniformly.

Concerning now the second quantity in the first supremum,

$$|a_n| \left| \int_0^B e^{-s}\, r_n(s)\, \Psi\big(g_n(s)\big|\theta\big)\; ds \right| \;,$$

we use again the fact that $|r_n(s)| \le \mathrm{K}\, e^{(\gamma^\star + \eta)s}$ and $\sup_{\theta \in \mathcal{V}^\star,\, x \in [0,A]} |\Psi(x|\theta)| < \infty$ in order to conclude.

The second supremum, related to

$$\left| \int_B^{B_n} e^{-s} \Psi\big(g_n(s) \big| \theta\big)\, g_n'(s)\ ds \right| ,$$

can be bounded by

$$\int_B^{B_n} e^{-s} \Big| \Psi\big(C_1\, e^{(\gamma^\star - \eta)s} \,\big|\, \theta\big) \Big|\, C_2\, e^{(\gamma^\star + \eta)s}\ ds \ = \ \mathcal{O}\bigg(\int_B^{B_n} e^{-(1-2\eta)s}\, ds \bigg) ,$$

where C_1 and C_2 are two constants. This last equality comes from the fact that B and B_n are large, then $\Psi(\cdot|\theta)$ is decreasing and $\Psi\big(C_1\, e^{(\gamma^\star-\eta)s} \,\big|\, \theta\big)$ is of order $\mathcal{O}\big(e^{-(\gamma^\star-\eta)s}\big)$ uniformly, which achieves the proof.

4.2. Conditions of integrability for $\bar{F}^*_{u_n}$

We would like to show that (C$_3$) is satisfied by $\bar{F}^*_{u_n}$. With this aim, let $e^{-s} = \bar{F}^*_{u_n}(y)$ which implies that $y = \bar{F}^{*\,-1}_{u_n}(e^{-s}) = g_n(s)$. Using the Potter bounds, we obtain that

$$\mathrm{cst}\, e^{(\gamma^\star - \eta)s} \ \le\ \bar{F}^{*\,-1}_{u_n}(e^{-s}) \ \le\ \mathrm{cst}\, e^{(\gamma^\star + \eta)s} \ .$$

Since $\bar{F}^*_{u_n}$ is decreasing, we deduce that

$$\bar{F}^*_{u_n}\big(\mathrm{cst}\, e^{(\gamma^\star + \eta)s}\big) \ \le\ e^{-s} \ \le\ \bar{F}^*_{u_n}\big(\mathrm{cst}\, e^{(\gamma^\star - \eta)s}\big) \ .$$

Let $y = \mathrm{cst}\, e^{(\gamma^\star \pm \eta)s}$ according to what we want to obtain. We have then $e^{-s} = (y/\mathrm{cst})^{-1/(\gamma^\star \pm \eta)}$, which implies that $\bar{F}^*_{u_n}(y) = \mathcal{O}(y^{-\beta})$ for some $\beta > 0$. The integrability condition (C$_3$) is then clearly satisfied by $\bar{F}^*_{u_n}$.

4.3. Hadamard differentiability of $\bar{F}^*_{u_n}$

Our aim in this appendix is to deal with

$$\int_0^\infty \frac{\bar{F}^*_{u_n}(y) - \bar{G}(y)}{a_n}\, \psi_{\theta^\star}(y)\ dy \ =$$

$$= \int_0^\infty e^{-s}\, \psi_{\theta^\star}(g_n(s))\, r_n(s)\ ds \ + \int_0^\infty e^{-(1-\gamma^\star)s}\, \frac{\psi_{\theta^\star}(g_n(s)) - \psi_{\theta^\star}(g(s))}{a_n}\ ds$$

with $|r_n(s)| \le \mathrm{cst}\, e^{(\gamma^\star + \eta)s}$.

We will use the fact that all the functions of the matrix $\psi_{\theta^\star}(x)$ are continuous and of order $\mathcal{O}\big(\frac{1}{x}\big)$ as $x \to \infty$. Denoting by ϕ such a function, we will first prove that

$$\int_0^\infty e^{-s}\,\phi(g_n(s))\,r_n(s)\;ds \underset{n\to\infty}{\longrightarrow} \int_0^\infty e^{-(1-\gamma^\star)s}\int_0^s e^{\rho t}\,dt\;\;\phi(g(s))\;ds\;.$$

With this aim, we will split the integral into two parts: the first one from 0 to A and the second one from A to infinity. The first part does not pose any problem. Therefore, we will look at the second one.

Using the properties of ϕ, we have

$$\big|\phi(g_n(s))\big| \le \frac{\text{cst}}{g_n(s)} \quad \text{for } s \ge A \quad \text{and for all } n \text{ sufficiently large}\,.$$

Using the lower Potter bound for $g_n(s)$, we have:

$$g_n(s) \ge \text{cst}\,e^{(\gamma^\star-\eta)s}\;, \qquad \text{cst} > 0, \quad \text{for } s \ge A \;\text{ and for all } n \text{ sufficiently large}\,.$$

Therefore, for the same s and n, we have

$$\big|\phi(g_n(s))\big| \le \text{cst}\,e^{(-\gamma^\star+\eta)s}\;,$$

which implies the domination of the function in the integral by $\text{cst}\,e^{-s(1-2\eta)}$.

We thus obtain the desired convergence using the fact that ϕ is continuous and that $r_n(s) \to e^{\gamma^\star s}\big(\int_0^s e^{\rho t}\,dt\big)$ as $n \to \infty$.

Now we have to study

$$\int_0^\infty e^{-(1-\gamma^\star)s}\;\frac{\phi(g_n(s)) - \phi(g(s))}{a_n}\;ds\;.$$

We have

$$\begin{aligned}\phi(g_n(s)) - \phi(g(s)) \;&=\; \phi\Big(g(s) + a_n R_n(s)\Big) - \phi(g(s))\\ &=\; a_n\,\phi'\Big(g(s) + \xi_n(s)\,a_n\,R_n(s)\Big)R_n(s)\;, \qquad 0 \le \xi_n(s) \le 1\;.\end{aligned}$$

We can use the Potter bound for $g(s) + \xi_n(s)\,a_n R_n(s)$. Therefore

$$\Big|\phi'\Big(g(s) + \xi_n(s)\,a_n\,R_n(s)\Big)\Big| \le\; \text{cst}\,e^{(-2\gamma^\star+2\eta)s}\;.$$

Recall that for R_n, we have the following bound

$$|R_n(s)| \le\; \text{cst}\,e^{(\gamma^\star+\eta)s}\;.$$

By gathering these various results, we obtain that the function in the integral is bounded by $\operatorname{cst} e^{-s(1-3\eta)}$.

Moreover, ϕ' is continuous. Therefore, since $\forall\, s,\ g_n(s) \to g(s)$ as $n \to \infty$ and $g(s) + \xi_n(s)\, a_n R_n(s)$ is located between $g(s)$ and $g_n(s)$, we have

$$\phi'\Big(g(s) + \xi_n(s)\, a_n\, R_n(s)\Big) \;\longrightarrow\; \phi'(g(s)) \;.$$

Finally, $R_n(s) \to \int_0^s e^{\gamma^\star u} \int_0^u e^{\rho t}\, dt\, du$.

By gathering the two limiting integrals, we obtain

$$\int_0^\infty e^{-(1-\gamma^\star)s} \Big(\int_0^s e^{\rho t}\, dt\Big)\, \phi(g(s))\; ds \;+ \\ + \int_0^\infty e^{-(1-\gamma^\star)s} \Big(\int_0^s e^{\gamma^\star z} \int_0^z e^{\rho t}\; dt\, dz\Big)\, \phi'(g(s))\; ds \;=:\; A + B \;.$$

By integrating by parts, we derive the following limit

$$B \;=\; -\int_0^\infty \phi(g(s))\, e^{-s} \Big(-\int_0^s e^{\gamma^\star z} \int_0^z e^{\rho t}\; dt\, dz \;+\; e^{\gamma^\star s} \int_0^s e^{\rho t}\, dt\Big)\; ds \;,$$

which implies that

$$A + B \;=\int_0^\infty \phi(g(s))\, e^{-s} \int_0^s e^{\gamma^\star z} \int_0^z e^{\rho t}\; dt\, dz\, ds \;.$$

Now, using the notations of the Introduction, it follows that

$$\begin{aligned} A + B \;&=\; \int_0^\infty \phi(g(s))\, e^{-s} I_{\gamma^\star,\rho}(s)\; ds \\ &=\; \int_0^\infty \phi(g(s))\, e^{\gamma^\star s}\, C_{\gamma^\star,\rho}(s)\; ds \\ &=\; \int_0^\infty \phi(y)\, D_{\gamma^\star,\rho}(y)\; dy \;, \end{aligned}$$

by the change of variable $y = g(s)$.

4.4. Existence and almost surely unicity of $\hat\theta_n$ for $\bar F^*_{u_n} + \frac{1}{\sqrt{k_n}}\, \alpha_{k_n}\big(\bar F^*_{u_n}\big)$

According to the proof for $\bar F^*_{u_n}$, it is sufficient to establish that

$$\frac{1}{\sqrt{k_n}} \int_0^\infty \alpha_{k_n}\big(\bar F^*_{u_n}(x)\big)\, D\psi_\theta(x)\; dx \;\underset{n\to\infty}{\longrightarrow}\; 0 \qquad \text{a.s.}\;,$$

uniformly in $\theta \in \mathcal{V}^\star$. We use the fact that $D\psi_\theta(x)$ is bounded uniformly in $\theta \in \mathcal{V}^\star$. We split the integral, as in the proof for $\bar{F}^*_{u_n}$, into two integrals, one from 0 to A and the other one from A to infinity.

Concerning the integral from A to infinity, we use the classical change of variable, leading to (see (3.9))

$$\frac{1}{\sqrt{k_n}} \int_{B_n}^{\infty} \left|\alpha_{k_n}(\mathrm{e}^{-s})\right| \frac{g_n'(s)}{g_n(s)}\, ds \ .$$

We use Mason's theorem (1981) (see Shorack and Wellner, 1986, p. 425) which implies that for a small $\varepsilon > 0$, we have

$$\frac{1}{\sqrt{k_n}} \left|\alpha_{k_n}(\mathrm{e}^{-s})\right| \ = \ \frac{(\ln k_n)^{\frac{1}{2}+\varepsilon}}{\sqrt{k_n}}\ \mathcal{O}\big(\mathrm{e}^{-\frac{s}{2}}\big) \qquad \text{a.s.}$$

when $s \to \infty$, as $k_n \to \infty$. To conclude, we apply to the term $\mathcal{O}\big(\mathrm{e}^{-\frac{s}{2}}\big)$ an integration by parts technique similar to the one used in (3.10).

Concerning the integral from 0 to A, Mason's theorem (1981) implies that

$$\frac{1}{\sqrt{k_n}} \left|\alpha_{k_n}\big(\bar{F}^*_{u_n}(x)\big)\right| \ = \ \frac{(\ln k_n)^{\frac{1}{2}+\varepsilon}}{\sqrt{k_n}}\ \mathcal{O}(1) \qquad \text{a.s.}$$

for $0 \le x \le A$, as $n \to \infty$.

4.5. Hadamard differentiability of $\bar{F}^*_{u_n} + \frac{1}{\sqrt{k_n}}\, \alpha_{k_n}\big(\bar{F}^*_{u_n}\big)$

We must check that the condition of integrability (C_3) is satisfied. We showed that $\bar{F}^*_{u_n}(y) = \mathcal{O}(y^{-\beta})$ when $y \to \infty$, uniformly for n sufficiently large, and for a constant $\beta > 0$. We thus deduce, via Mason (1981), that

$$\frac{1}{\sqrt{k_n}} \left|\alpha_{k_n}\big(\bar{F}^*_{u_n}(y)\big)\right| \ = \ \frac{(\ln k_n)^{\frac{1}{2}+\varepsilon}}{\sqrt{k_n}}\ O\big(y^{-\frac{\beta}{2}}\big) \qquad \text{a.s.}$$

as $y \to \infty$, uniformly for n sufficiently large.

Now, we will show that there exists some versions $\widetilde{\alpha}_{k_n}$ of α_{k_n} and a brownian bridge $\mathbb{B}$ on $[0,1]$ such that

$$\int_0^{\infty} \widetilde{\alpha}_{k_n}\big(\bar{F}^*_{u_n}(x)\big)\, \psi_{\theta^\star}(x)\ dx \ = \ \int_0^{\infty} \mathbb{B}\big(\bar{G}(x)\big)\, \psi_{\theta^\star}(x)\ dx \ + \ o_P(1)$$

as $n \to \infty$. With this aim, we will use the fact that $|\psi_{\theta^\star}(x)| = \mathcal{O}(1/x)$ as $x \to \infty$, and the invariance principle for the weighted empirical process given in Einmahl and Mason (1992). We again split the initial integral into an integral from 0 to A and an integral from A to infinity, and we carry out the usual change of variable.

We start with $\int_{B_n}^{\infty} \alpha_{k_n}(\mathrm{e}^{-s})\, \psi_{\theta^\star}(g_n(s))\, g_n'(s)\, ds$. Since $|\psi_{\theta^\star}(x)| = \mathcal{O}(1/x)$ as $x \to \infty$, this integral is of order $\mathcal{O}\Big(\int_{B_n}^{\infty} |\alpha_{k_n}(\mathrm{e}^{-s})|\, \frac{g_n'(s)}{g_n(s)}\, ds\Big)$. Here, we will change α_{k_n} into $\widetilde{\alpha}_{k_n}$ (Einmahl and Mason, 1992) for $B_n \le s \le \ln k_n$, with an error term of order $\mathcal{O}_P\big(k_n^{-\nu}\, \mathrm{e}^{-(\frac{1}{2}-\nu)s}\big)$ for $0 \le \nu < 1/4$. We obtain therefore

$$\begin{aligned}
&\int_{B_n}^{\ln k_n} \widetilde{\alpha}_{k_n}(\mathrm{e}^{-s})\, \psi_{\theta^\star}(g_n(s))\, g_n'(s)\; ds \;= \\
&\qquad = \int_{B_n}^{\ln k_n} \mathbb{B}(\mathrm{e}^{-s})\, \psi_{\theta^\star}(g_n(s))\, g_n'(s)\; ds \\
&\qquad\quad + \mathcal{O}_P\left(k_n^{-\nu} \int_{B_n}^{\ln k_n} \mathrm{e}^{-(\frac{1}{2}-\nu)s}\, \big|\psi_{\theta^\star}(g_n(s))\big|\, g_n'(s)\; ds\right)
\end{aligned} \tag{4.4}$$

as $n \to \infty$. The error term is

$$\mathcal{O}_P\left(k_n^{-\nu} \int_{B_n}^{\ln k_n} \mathrm{e}^{-(\frac{1}{2}-\nu)s}\, \frac{g_n'(s)}{g_n(s)}\; ds\right) ,$$

and we conclude that it tends to 0 by integrating by parts.

Now, we will study the integral from $\ln k_n$ to infinity by using again the fact that $|\psi_{\theta^\star}(x)| = \mathcal{O}(1/x)$ as $x \to \infty$. According to Jaeschke's theorem (see Shorack and Wellner, 1986, p. 600), this integral is of order

$$\sqrt{\ln\ln k_n}\; \mathcal{O}_P\left(\int_{B_n}^{\infty} \mathrm{e}^{-s/2}\, \frac{g_n'(s)}{g_n(s)}\; ds\right) , \tag{4.5}$$

and again by integrating by parts, the result follows.

Combining (4.1) with (4.2), we obtain that

$$\int_{B_n}^{\infty} \widetilde{\alpha}_{k_n}(\mathrm{e}^{-s})\, \psi_{\theta^\star}(g_n(s))\, g_n'(s)\; ds \;=\; \int_{B_n}^{\ln k_n} \mathbb{B}(\mathrm{e}^{-s})\, \psi_{\theta^\star}(g_n(s))\, g_n'(s)\; ds \;+\; o_P(1)$$

as $n \to \infty$.

Therefore, we only have to study $\int_0^{B_n} \alpha_{k_n}(\mathrm{e}^{-s})\, \psi_{\theta^\star}(g_n(s))\, g_n'(s)\, ds$. Again, we split into two integrals, one from 0 to $-\ln(1-\frac{1}{k_n})$ and the other from $-\ln(1-\frac{1}{k_n})$ to B_n. The first one is clearly $o_P(1)$. For the second one, we use the fact that $\psi_{\theta^\star}(x)$ is bounded for $0 \le x \le A$, and we replace $\widetilde{\alpha}_{k_n}(\mathrm{e}^{-s})$ by $\mathbb{B}(\mathrm{e}^{-s})$, which leads to an error term of order

$$\mathcal{O}_P\left(k_n^{-\nu} \int_{-\ln\left(1-\frac{1}{k_n}\right)}^{B_n} \mathrm{e}^{-(\frac{1}{2}-\nu)s}\, g_n'(s)\; ds\right)$$

as $n \to \infty$. Remark now that $(B_n)_{n\ge 1}$ is bounded. Indeed, the Potter bounds imply that $0 \le g_n(s) \le \mathrm{cst}\, e^{(\gamma^\star+\eta)s}$. If we note $s = g_n^{-1}(y)$, we have $0 \le y \le \mathrm{cst}\, e^{(\gamma^\star+\eta)s}$, so $s \le \frac{\ln(y/\mathrm{cst})}{\gamma^\star+\eta}$. This leads to $B_n \ge \frac{\ln(A/\mathrm{cst})}{\gamma^\star+\eta}$, with a similar result in the other side.

Consequently, using again the Potter bounds, but this time for $g_n'(s)$, and the fact that $(B_n)_{n\geq 1}$ is bounded, the preceding error term is $o_P(1)$.

Finally, we obtain the following result

$$\int_0^\infty \widetilde{\alpha}_{k_n}(\mathrm{e}^{-s})\,\psi_{\theta^\star}(g_n(s))\,g_n'(s)\;ds \;=\; = \int_{-\ln\left(1-\frac{1}{k_n}\right)}^{\ln k_n} \mathbb{B}(\mathrm{e}^{-s})\,\psi_{\theta^\star}(g_n(s))\,g_n'(s)\;ds \;+\; o_P(1)\;.$$

Now, we only have to let n tending to infinity in the integral of the second member. Since $\mathbb{B}$ is a.s. continuous, $\ln k_n \to \infty$, $-\ln(1-\frac{1}{k_n}) \to 0$, $g_n(s) \to g(s)$ and $g_n'(s) \to g'(s)$ for all s as $n \to \infty$, we only have to establish an a.s. domination for $\mathbb{B}(\mathrm{e}^{-s})\,|\psi_{\theta^\star}(g_n(s))|\,g_n'(s)$ in order to apply the Lebesgue dominated convergence theorem. We use again the fact that $|\psi_{\theta^\star}(g_n(s))| = \mathcal{O}(1/g_n(s))$ as $s \to \infty$ for n sufficiently large, and we conclude using the Potter bounds on $g_n(s)$ and $g_n'(s)$, and the law of iterated logarithm for $\mathbb{B}(t)$, $t \to 0$.

ACKNOWLEDGMENTS

We thank the referees for valuable comments.

REFERENCES

[1] Balkema, A. and de Haan, L. (1974). Residual life time at a great age, *Annals of Probability*, **2**, 792–801.

[2] Beirlant, J.; Dierckx, G.; Goegebeur, Y. and Matthys, G. (1999). Tail index estimation and an exponential regression model, *Extremes*, **2**(2), 177–200.

[3] Bingham, N.H.; Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation*, Cambridge, U.K.: Cambridge University Press.

[4] Cohen, J.P. (1982). Convergence rates for the ultimate and penultimate approximations in extreme-value theory, *Advances in Applied Probability*, **14**, 833–854.

[5] Dekkers, A.L.M. and de Haan, L. (1989). On the estimation of the extreme-value index and large quantile estimation, *Annals of Statistics*, **17**, 1795–1832.

[6] Diebolt, J.; Guillou, A. and Worms, R. (2003). Asymptotic behaviour of the probability-weighted moments and penultimate approximation, *ESAIM, Probability and Statistics*, **7**, 219–238.

[7] Drees, H. (1998). On smooth statistical tail functionals, *Scandinavian Journal of Statistics*, **25**(1), 187–210.

[8] Drees, H.; Ferreira, A. and de Haan, L. (2004). On maximum likelihood estimation of the extreme value index, *Annals of Applied Probability*, **14**(3), 1179–1201.

[9] Einmahl, U. and Mason, D.M. (1992). Approximation to permutation and exchangeable processes, *Journal of Theoretical Probability*, **5**, 101–126.

[10] Galambos, J. (1978). *Asymptotic Theory of Extreme Order Statistics*, Krieger, Malabar, Florida.

[11] Gnedenko, B.V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire, *Annals of Mathematics*, **44**, 423–453.

[12] Gomes, M.I. (1984). Penultimate limiting forms in extreme value theory, *Annals of the Institute of Statistical Mathematics*, **36**, 71–85.

[13] Gomes, I. and de Haan, L. (2000). Approximation by penultimate extreme value distributions, *Extremes*, **2**, 71–85.

[14] Gomes, M.I. and Pestana, D.D. (1987). *Non standard domains of attraction and rates of convergence*, In "New perspectives in theoretical and applied statistics", (Wiley Ser. Probab. Math. Statist. Probab. Math. Statist.), Wiley, New York, 467–477.

[15] Hall, P. (1982). On some simple estimates of an exponent of regular variation, *Journal of the Royal Statistical Society Serie B*, **44**, 37–42.

[16] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution, *Annals of Statistics*, **3**, 1163–1174.

[17] Hosking, J. and Wallis, J. (1987). Parameter and quantile estimation for the generalized Pareto distribution, *Technometrics*, **29**, 339–349.

[18] Leadbetter, M.R.; Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*, Springer, Berlin.

[19] Mason, D.M. (1981). Bounds for weighted empirical distribution functions, *Annals of Probability*, **9**, 881–884.

[20] Pickands III, J. (1975). Statistical inference using extreme order statistics, *Annals of Statistics*, **3**, 119–131.

[21] Shorack, G.R. and Wellner, J.A. (1986). *Empirical Processes with Applications to Statistics*, Wiley New York.

[22] Smith, R.L. (1987). Estimating tails of probability distributions, *Annals of Statistics*, **15**, 1174–1207.

[23] Worms, R. (2000). *Vitesses de convergence pour l'approximation des queues de distributions*, Thèse de doctorat de l'Université de Marne-la-Vallée.

[24] Worms, R. (2002). Penultimate approximation for the distribution of the excesses, *ESAIM, Probability and Statistics*, **6**, 21–31.